Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 363 (2003) 35–41
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Minimal representations of inverted Sylvester and Lyapunov operators

Tobias Damm

Zentrum Mathematik, TU München, München 80333, Germany

Received 23 March 2001; accepted 14 November 2001

Submitted by R. Nabben

Abstract

We derive minimal representations for the inverses of Lyapunov and Sylvester operators.

1. Introduction

Let $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$ and define the Sylvester operator $\mathscr{S}_{A,B} : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ by

$$\mathscr{S}_{A,B}(X) = AX - XB.$$

If $n = m$ and $B = -A^*$, then $\mathscr{S}_{A,B}$ is a Lyapunov-operator, and we write $\mathscr{S}_{A,B} = \mathscr{L}_A$. It is assumed that $\mathscr{S}_{A,B}$ is nonsingular, i.e., $0 \notin \sigma(A) - \sigma(B)$.

During the Oberwolfach meeting on ‘Nonnegative matrices, M-matrices and their generalizations’, V. Mehrmann raised a question about minimal representations for the inverse of the Lyapunov-operator:

Clearly, there exist matrices $V_i \in \mathbb{C}^{m\times m}$, $W_i \in \mathbb{C}^{n\times n}$ such that for all $Y \in \mathbb{C}^{m\times n}$ we have

$$\mathscr{L}_A^{-1}(Y) = \sum_{i=1}^{N} V_i Y W_i. \tag{1}$$

But what is the minimal number N of terms in this sum?

E-mail address: damm@ma.tum.de (T. Damm).

0024-3795/03/$ - see front matter
PII: S0024-3795(01)00601-2

In the present paper we show that this question can be answered straightforwardly by combining results from [4] and [3]. More precisely, our aim is to prove the following:

Theorem 1.1. *Let* $\nu_A = \deg \mu_A$ *and* $\nu_B = \deg \mu_B$ *denote the degrees of the minimal polynomials of A and B*, *respectively*, *and* $\nu = \min\{\nu_A, \nu_B\}$.

(i) *The inverse* $\mathscr{S}_{A,B}^{-1} : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ *has a representation of the form*

$$\mathscr{S}_{A,B}^{-1}(Y) = \sum_{i=1}^{\nu} V_i Y W_i, \quad V_i \in \mathbb{C}^{n\times n}, \quad W_i \in \mathbb{C}^{n\times n}.$$

(ii) *If* $B = -A^*$, *i.e.* $\mathscr{S}_{A,B} = \mathscr{L}_A$, *there exists a symmetric representation*

$$\mathscr{L}_A^{-1}(Y) = \sum_{i=1}^{\nu_A} \varepsilon_i A_i Y A_i^*, \quad \varepsilon_i = \pm 1, \quad A_i \in \mathbb{C}^{n\times n}$$

with $\varepsilon_i = 1$ *for all i if and only if* $\sigma(A) \subset \mathbb{C}_+$.

(iii) *If an arbitrary representation*

$$\mathscr{S}_{A,B}^{-1}(Y) = \sum_{i=1}^{N} V_i Y W_i, \quad V_i \in \mathbb{C}^{m\times m}, \quad W_i \in \mathbb{C}^{n\times n}$$

is given, *then necessarily* $N \geqslant \nu$.

Remark 1.2

(a) If $\varepsilon_i = 1$ for all i in (ii), then $\mathscr{L}_A^{-1}$ is called *completely positive* (compare [2]). Another approach to verify complete positivity of $\mathscr{L}_A^{-1}$ can be based on [6], or [1].

(b) From our proof (and Theorem 3.2) it is immediate to see that an analogous result holds for the discrete Sylvester operator $\mathscr{S}_{A,B} : X \mapsto X - AXB$. In this case, of course, $\mathscr{S}_{A,B}$ is invertible if and only if $1 \notin \sigma(A)\sigma(B)$, and $\mathscr{S}_{A,B}$ specializes to the discrete Lyapunov operator if $B = A^*$. Taking this into account, all assertions of Theorem 1.1 carry over literally.

(c) In our proof of Theorem 1.1 we make use of Theorem 4 in [3]. In fact one can easily use Theorems 1 and 4 in [3] to prove a stronger version of (i), where $V_i = p_i(A)$ and $W_i = q_i(B)$ with polynomials p_i and q_i. Our main concern is, however, that one needs *at least* ν terms in the representation.

We begin with some facts from representation theory for mappings between matrix spaces. These results are not new, but we take the chance of giving a concise self-contained presentation, providing all the details we need.

2. Representation of mappings between matrix spaces

We consider linear mappings $\mathscr{T} : \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{p\times q}$ and try to give a representation of the form $\mathscr{T}(X) = \sum_{i=1}^{N} V_i X W_i$ with $V_i \in \mathbb{C}^{p\times m}$, $W_i \in \mathbb{C}^{n\times q}$.

Obviously $\mathscr{T}$ can also be regarded as a linear mapping between the vector spaces $\mathbb{C}^{mn}$ and $\mathbb{C}^{pq}$. To make use of this observation, we recall the definition and some basic properties of the Kronecker product and the vec-operator.

Definition 2.1. Let $V = (v_{ij}) = (v_1, \ldots, v_n) \in \mathbb{C}^{m\times n}$ and $U = (u_{ij}) \in \mathbb{C}^{p\times q}$. Then

$$V \otimes U = \big(v_{ij} U\big) \in \mathbb{C}^{mp\times nq}$$

and

$$\mathrm{vec} V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^{nm}.$$

Lemma 2.2 (e.g. [5]). *Let U_1, U_2, V_1, V_2, and X be matrices of appropriate sizes. Then*

$$(V_1 \otimes U_1)(V_2 \otimes U_2) = (V_1 V_2) \otimes (U_1 U_2)$$

and

$$\mathrm{vec}(U_1 X V_1) = (U_1^{\mathrm{T}} \otimes V_1)\mathrm{vec}\, X.$$

In the sequel let $e_i^{(m)}$ denote the ith canonical unit vector in $\mathbb{C}^m$ and $E_{ij}^{(mn)} = e_i^{(m)} e_j^{(n)*}$ the $m \times n$ matrix with the only nonzero entry 1 in the ith row and jth column. By E^{mn} we denote the $m^2 \times n^2$-block matrix

$$\big(E_{ij}^{(mn)}\big)_{i,j=1}^{m,n}$$

and with the mapping $\mathscr{T}$ we associate the $mp \times nq$ matrix

$$(I_{pq} \otimes \mathscr{T})(E^{(mn)}) = \Big(\mathscr{T}\big(E_{ij}^{(mn)}\big)\Big)_{i,j=1}^{(m,n)}.$$

The following simple identity plays a crucial role.

Lemma 2.3. *If $V = (v_1, \ldots, v_m) \in \mathbb{C}^{p\times m}$ and $W = (w_1, \ldots, w_q) \in \mathbb{C}^{n\times q}$, then*

$$\Big(V E_{ij}^{(mn)} W^*\Big)_{i,j=1}^{m,n} = \mathrm{vec} V (\mathrm{vec} W)^*. \tag{2}$$

Proof. By Lemma 2.2 we have

$$\begin{aligned}\operatorname{vec}\left(V E_{ij}^{(mn)} W^*\right) &= \bar{W} \otimes V \operatorname{vec} E_{ij}^{(mn)} \\ &= \left(\bar{w}_1 \otimes v_1, \bar{w}_1 \otimes v_2, \ldots, \bar{w}_q \otimes v_m\right) e_{(j-1)n+i}^{(nm)} \\ &= \bar{w}_j \otimes v_i = \operatorname{vec} v_i w_j^*.\end{aligned}$$

Thus

$$V E_{ij}^{(mn)} W^* = v_i w_j^*$$

and

$$\left(V E_{ij}^{(mn)} W^*\right)_{i,j=1}^{m,n} = \left(v_i w_j^*\right)_{i,j=1}^{m,n} = \operatorname{vec} V (\operatorname{vec} W)^*. \qquad \square$$

Theorem 2.4. *Let $\mathscr{T} : \mathbb{C}^{m\times n} \longrightarrow \mathbb{C}^{p\times q}$ be linear. The matrices $V_1, \ldots, V_N \in \mathbb{C}^{p\times m}$, $W_1, \ldots, W_N \in \mathbb{C}^{n\times q}$ yield a representation of $\mathscr{T}$:*

$$\forall X \in \mathbb{C}^{m\times n}: \quad \mathscr{T}(X) = \sum_{i=1}^{N} V_i X W_i \tag{3}$$

if and only if

$$(I_{pq} \otimes \mathscr{T})(E^{(mn)}) = \sum_{i=1}^{N} \operatorname{vec} V_i (\operatorname{vec} W_i)^*. \tag{4}$$

In particular, the minimal number of summands is $\nu = \operatorname{rk}(I_{pq} \otimes \mathscr{T})(E^{(mn)})$, and one can choose a representation with matrices $V_1, \ldots, V_\nu \in \mathbb{C}^{p\times m}$, $W_1, \ldots, W_\nu \in \mathbb{C}^{n\times q}$ such that both sets $\{\operatorname{vec} V_1, \ldots, \operatorname{vec} V_\nu\}$ and $\{\operatorname{vec} W_1, \ldots, \operatorname{vec} W_\nu\}$ are orthogonal.

Proof. By Lemma 2.3, identity (3) clearly implies (4); vice versa (4) implies $\mathscr{T}(X) = \sum_{i=1}^{N} V_i X W_i$ for all $X = E_{ij}^{mn}$, and thus for all $X \in \mathbb{C}^{m\times n}$.

A representation with a minimal number of summands is given, e.g., by a singular value decomposition of $(I_{pq} \otimes \mathscr{T})(E^{(mn)})$, in which case also the orthogonality property holds. $\square$

We now consider mappings between spaces of quadratic matrices.

Definition 2.5. Let $\mathscr{T} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{m\times m}$ be a linear map. Then $\mathscr{T}$ is called *Hermitian-preserving* if $\mathscr{T}\left(\mathscr{H}^n\right) \subset \mathscr{H}^m$.

It is immediate to see that the Lyapunov operator is Hermitian preserving. The following representation result can be found in [4]; the proof is adapted from [2].

Theorem 2.6. *For a linear map $\mathscr{T} : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{m\times m}$ the following are equivalent*:

1. *$\mathscr{T}$ is Hermitian-preserving.*
2. $\forall X \in \mathbb{C}^{n\times n}: \mathscr{T}(X^*) = (\mathscr{T}(X))^*$.

3. *The* $nm \times nm$ *matrix* $(I_{mm} \otimes \mathscr{T})(E^{(nn)}) = \left(\mathscr{T}\left(E_{ij}^{(nn)}\right)\right)_{i,j=1}^{n}$ *is Hermitian.*
4. *There exist matrices* $V_1, \ldots, V_N \in \mathbb{C}^{m\times n}$ *and numbers* $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ *such that* $\forall X \in \mathbb{C}^{n\times n} : \mathscr{T}(X) = \sum_{i=1}^{N} \varepsilon_i V_i X V_i^*$.
   In particular one can choose $N = \operatorname{rk}(I_{mm} \otimes \mathscr{T})(E^{(nn)})$.
   One can choose $\varepsilon_i = 1$ *for all* i *if and only if* $(I_{mm} \otimes \mathscr{T})(E^{(nn)}) \geqslant 0$.

Proof. $1 \Rightarrow 2$. For skew-Hermitian matrices S (i.e., iS Hermitian) we have on the one hand

$$\mathscr{T}(\mathrm{i}S) = \mathscr{T}(\mathrm{i}S)^* = -\mathrm{i}\mathscr{T}(S)^*$$

and on the other

$$\mathscr{T}(\mathrm{i}S) = \mathscr{T}((\mathrm{i}S)^*) = \mathscr{T}(-\mathrm{i}S^*) = -\mathrm{i}\mathscr{T}(S^*),$$

which yields $\mathscr{T}(S)^* = \mathscr{T}(S^*)$.

As any matrix A can be decomposed in $A = H + S$, where $H = \frac{1}{2}(A + A^*)$ is Hermitian and $S = \frac{1}{2}(A - A^*)$ is skew-Hermitian, the result follows.

$2 \Rightarrow 3$. This follows from

$$\begin{aligned}\left(\left(\mathscr{T}\left(E_{ij}^{(nn)}\right)\right)_{i,j=1}^{n}\right)^* &= \left(\left(\mathscr{T}\left(E_{ij}^{(nn)}\right)^*\right)_{j,i=1}^{n}\right)\\ &= \left(\left(\mathscr{T}\left(E_{ij}^{(nn)*}\right)\right)_{j,i=1}^{n}\right)\\ &= \left(\left(\mathscr{T}\left(E_{ji}^{(nn)}\right)\right)_{j,i=1}^{n}\right)\\ &= \left(\left(\mathscr{T}\left(E_{ij}^{(nn)}\right)\right)_{i,j=1}^{n}\right).\end{aligned}$$

$3 \Rightarrow 4$. By Sylvester's Theorem there is a decomposition

$$\left(\mathscr{T}\left(E_{ij}^{(nn)}\right)\right)_{i,j=1}^{n} = \sum_{l=1}^{N} \varepsilon_l \operatorname{vec}\bar{V}_l (\operatorname{vec}\bar{V}_l)^* \tag{5}$$

for appropriate V_l, where N is the rank of the matrix on the left. By Eq. (2) we can also write

$$\left(\mathscr{T}\left(E_{ij}^{(nn)}\right)\right)_{i,j=1}^{n} = \left(\sum_{l=1}^{N} \varepsilon_l V_l E_{ij}^{(nn)} V_l^*\right)_{i,j=1}^{n}$$

that means

$$\forall E_{ij}^{(nn)} : \quad \mathscr{T}\left(E_{ij}^{(nn)}\right) = \sum_{l=1}^{N} \varepsilon_l V_l E_{ij}^{(nn)} V_l^*$$

and thus the same holds with $E_{ij}^{(nn)}$ replaced by any $X \in \operatorname{span}\{E_{ij}^{(nn)} \mid i, j = 1, \ldots, n\} = \mathbb{C}^{n\times n}$.

Moreover it follows from Sylvester's Theorem that one can choose $\varepsilon_i = 1$ for all i if and only if $(I_{mm} \otimes \mathscr{T})(E^{(nn)}) \geqslant 0$.

$4 \Rightarrow 1$. Since for each l the matrix $V_l^* X V_l$ is Hermitian if X is Hermitian the sum is Hermitian, too. $\square$

3. Proof of Theorem 1.1

In view of Theorems 2.4 and 2.6 we need to determine the rank of the matrix

$$\mathscr{X} = (I_{mn} \otimes \mathscr{S}_{A,B}^{-1})(E^{(mn)}) = \left(\mathscr{S}_{A,B}^{-1}\big(E_{ij}^{(mn)}\big)\right)_{i,j=1}^{m,n}. \tag{6}$$

To this end we apply the following result from [3, Theorem 4]. For arbitrary $k \in \mathbb{N}$ and a given matrix vector pair $(C, x) \in \mathbb{C}^{k\times k} \times \mathbb{C}^k$ let $K(C, x) = (x, Cx, \ldots, C^{k-1}x)$ denote the reachability matrix.

Theorem 3.1. *Let $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, $v \in \mathbb{C}^m$, $w \in \mathbb{C}^n$, and assume that the Sylvester equation $AX - XB = vw^*$ has a unique solution X. Then*

$$\operatorname{rk} X = \min\left\{\operatorname{rk}(A, v), \operatorname{rk}(B^* w)\right\}.$$

The analogous result for the discrete Sylvester equation was proven in [7].

Theorem 3.2. *Let $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, $v \in \mathbb{C}^m$, $w \in \mathbb{C}^n$, and assume that the Sylvester equation $X - AXB = vw^*$ has a unique solution X. Then*

$$\operatorname{rk} X = \min\left\{\operatorname{rk}(A, v), \operatorname{rk}(B^* w)\right\}.$$

If in (6) we partition $\mathscr{X} = \left(\mathscr{X}_{ij}\right)_{i,j=1}^{m,n}$ conformably with $E^{(mn)}$ such that $\mathscr{X}_{ij} \in \mathbb{C}^{m\times n}$, then we can also write $A\mathscr{X}_{ij} - \mathscr{X}_{ij}B = E_{ij}^{(mn)}$ and thus

$$(I_n \otimes A)\mathscr{X} - \mathscr{X}(I_m \otimes B) = E^{(mn)}. \tag{7}$$

In other words $\mathscr{X}$ is a solution of the Sylvester equation (7) with the right-hand side $E^{(mn)}$. By (2) (with $V = I_m$, $W = I_n$), we have $E^{(mn)} = \operatorname{vec} I_m (\operatorname{vec} I_n)^*$.

Let now $k = m$ or $k = n$ such that $C \in \mathbb{C}^{k\times k}$. It follows from Theorem 3.1 that

$$\begin{aligned}\operatorname{rk} \mathscr{X} &= \min_{C\in\{A,B^*\}} \operatorname{rk}\left(\operatorname{vec} I_k, (I_k \otimes C)\operatorname{vec} I_k, \ldots, (I_k \otimes C)^{k^2}\operatorname{vec} I_k\right)\\ &= \min_{C\in\{A,B^*\}} \operatorname{rk}\begin{pmatrix} e_1^{(k)} & Ce_1^{(k)} & \ldots & C^{k^2}e_1^{(k)}\\ e_2^{(k)} & Ce_2^{(k)} & \ldots & C^{k^2}e_2^{(k)}\\ \vdots & \vdots & \ddots & \vdots\\ e_k^{(k)} & Ce_k^{(k)} & \ldots & C^{k^2}e_k^{(k)} \end{pmatrix}.\end{aligned}$$

It is clear that the first $\nu_C + 1$ columns are linearly dependent, if $\nu_C = \deg \mu_C$ denotes the degree of the minimal polynomial of C; hence $\operatorname{rk} \mathscr{X} \leqslant \nu_A$ and $\operatorname{rk} \mathscr{X} \leqslant \nu_B$.

If we consider arbitrary linear combinations of the block rows, we find

$$\operatorname{rk} \mathscr{X} \geqslant \min_{C \in \{A, B^*\}} \max_{x \in \mathbb{C}^k} \operatorname{rk}\big(x, \ldots, Cx, \ldots, C^k x\big) = \min_{C \in \{A, B^*\}} \nu_C = \nu.$$

Hence $\operatorname{rk}\mathscr{X} = \min\{\nu_A, \nu_B\}$, which proves (i), (iii) and the first part of (ii).

It remains to show that the ε_i can be chosen positive if and only if $\sigma(A) \in \mathbb{C}_+$. If all ε_i are positive, then $\mathscr{L}_A^{-1}$ is positive, i.e., $\mathscr{L}_A^{-1}(Y) \geqslant 0$ for all $Y \geqslant 0$ (the converse, however, is not true, see [2]). But it is well known (e.g. [5]) that $\mathscr{L}_A^{-1}$ is positive if and only if $\sigma(A) \subset \mathbb{C}_+$. Hence $\varepsilon_i \geqslant 0$ for all i implies $\sigma(A) \subset \mathbb{C}_+$.

To prove the converse, we need to show $\mathscr{X} \geqslant 0$. Let $\mathscr{L}_A = R_+ - R_-$ with

$$R_\pm(X) = \frac{1}{2}(A \pm I)X(A \pm I)^*.$$

One readily verifies that $\rho(R_+^{-1}R_-) < 1$ if $\sigma(A) \subset \mathbb{C}_+$, whence

$$\mathscr{L}_A^{-1}(Y) = \sum_{k=0}^{\infty} R_+^{-(k+1)} \circ R_-^k(Y) = \sum_{k=0}^{\infty} V_k Y V_k^*$$

with $V_k = (A + I)^{-(k+1)}(A - I)^k$.

Therefore $\mathscr{X} = \sum_{k=0}^{\infty}(I_n \otimes V_k)E^{(nn)}(I_n \otimes V_k)^* \geqslant 0$, which completes the proof. $\square$

The analogous result for the discrete Sylvester operator $\mathscr{S}_{A,B}(X) = X - AXB$ follows if one applies Theorem 3.2 instead of Theorem 3.1.

References

[1] R. Bhatia, L. Elsner, Sensitivity of the solutions of some matrix equations, preprint 97-106, SFB 343, Universität Bielefeld, 1997.

[2] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. 10 (1975) 285–290.

[3] E. de Souza, S. Bhattacharyya, Controllability observability and the solution of $AX - XB = C$, Linear Algebra Appl. 39 (1981) 167–188.

[4] R.D. Hill, Linear transformations which preserve Hermitian matrices, Linear Algebra Appl. 6 (1973) 257–262.

[5] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[6] P. Lancaster, Explicit solutions of linear matrix equations, SIAM Rev. 12 (4) (1970) 544–566.

[7] H.K. Wimmer, Linear matrix equations, controllability and observability, and the rank of solutions, SIAM J. Matrix Anal. Appl. 9 (1988) 570–578.